(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:

19.03.2003 Bulletin 2003/12

(21) Application number: 01830589.6

(22) Date of filing: 14.09.2001

(84) Designated Contracting States:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

Designated Extension States:

AL LT LV MK RO SI

(71) Applicant: STMicroelectronics S.r.l. 20041 Agrate Brianza (Milano) (IT)

(72) Inventors:

 Concion, Davide 36010 Monticello Conte Otto (Vincenza) (IT)

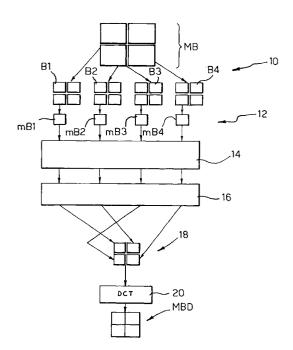
(51) Int Cl.7: **H04N 7/26**, G06T 9/00

- Piccinelli, Emiliano 20052 Monza (Milano) (IT)
- Pau, Danilo 20099 Sesto San Giovanni (Milano) (IT)
- (74) Representative: Bosotti, Luciano et al c/o Buzzi, Notaro & Antonielli d'Oulx Via Maria Vittoria 18 10123 Torino (IT)

(54) A process for the format conversion of MPEG bitstreams, a system and computer program product therefor

(57)The DCT macroblocks (MB) comprised in an MPEG video bitstream are divided into blocks (B1, B2, B3, B4), each of which comprises a plurality of microblocks. In each DCT block (B1, B2, B3, B4), the significant frequencies are identified and preserved, isolating a corresponding microblock (mB1, mB2, mB3, mB4) preferably consisting of the microblock on the top left of each block (B1, B2, B3, B4) and setting to zero the coefficients of the remaining microblocks. On the microblock thus isolated there is performed an inverse discrete cosine transform (16), and the microblock thus obtained is merged (18) with the homologous microblocks obtained from the other blocks comprised in a respective starting macroblock (MB), so as to give rise to a merging block. The merging block thus obtained undergoes a discrete cosine transform (20) so as to obtain a final block, which can be assembled into a macroblock (MBD) with converted format.





15

Description

Field of the invention

[0001] The present invention relates to the processing of bitstreams encoded according to the MPEG standard. [0002] The MPEG (Moving Pictures Experts Group) standard proposes a set of algorithms dedicated to the compression of sequences of digital (audio/video) signals. The techniques used are based on the reduction in spatial and temporal redundancy of the sequence.

[0003] According to the MPEG standard, reduction in spatial redundancy is obtained by independently compressing the individual images, using a discrete cosine transform (DCT), quantization and Huffman coding.

[0004] Reduction in temporal redundancy is obtained by exploiting the correlation that exists between successive and/or temporally close images in the sequence. Approximately, it is assumed that each portion of an image could be expressed locally as the translation of a portion of a previous and/or subsequent image in the sequence.

[0005] The foregoing corresponds to criteria that are well known to the prior art and hence are such as not to require further illustration herein.

Description of prior art

[0006] In the actual use of the MPEG standard it is possible to transmit (or record) films, or in general video sequences, on a variety of channels and media, each of which has its own characteristics of capacity, speed and cost.

[0007] The transmission of a film (or of a video program in general) starting from the master recording may be done, for instance, on DVD media, via satellite, radio antenna, cable, and — to an ever-increasing extent — via telematic networks, such as Internet, and/or to mobile terminals, such as UMTS mobile terminals.

[0008] In order to make available a wide range of the aforementioned transmission modalities and to expand it even further, it is necessary to be able to adapt the characteristics of the MPEG bitstream to the transmission/storage channel, as well as to the characteristics of the display units employed to display the corresponding images.

[0009] The foregoing entails, among other things, the need to convert the format of an MPEG video bitstream in such a way that the resulting sequence is reduced, for example, to one quarter of the original format.

[0010] In order to obtain the above result, it is possible, for example, to apply the "canonical" solution represented in Figure 1.

[0011] In the case of Figure 1, an input MPEG video bitstream, designated by IS, is entered into a decoder 100, which is able to read and decode the bitstream IS with a format corresponding to a number of pixels Hor x Vert per image.

[0012] The signal thus decoded is fed to a set of down-sampling filters 120 which are able to downsample the image starting from the original format Hor x Vert to a reduced format (Hor/N) x (Vert/M).

[0013] The signal thus subjected to conversion of format is entered into an encoder 140, which is able to generate a syntax in compliance with the MPEG standard, in view of its being sent on to a transmission/storage channel C and received in a decoder 160.

[0014] The decoder 160 is able to read and decode an MPEG bitstream with a number of pixels (Hor/N) x (Vert/M) per image, hence one with a reduced format, generating at output a corresponding video sequence OS.

[0015] The above technique in effect entails a conversion of the original MPEG bitstream (irrespective of whether this is in compliance with the MPEG2 specification or with the MPEG4 specification) into decoded images that are a sequence of frames. The change of resolution is based upon a filtering action — this being preferably a finite-impulse-response (FIR) filtering — which is able to bring about a conversion based upon the availability of a certain number of pixels for each component of luminance and chrominance of the image. The pixels in question are multiplied by appropriate weights, and the results are accumulated and divided by the sum of said weights.

[0016] Certain pixels are not transmitted in the resulting image, depending upon the factor of change of resolution chosen. For this reason, starting from a bitstream decoded with an arbitrary bitrate B1 it is always possible to obtain a bitstream of bitrate B2 by simply connecting the output of the decoder to the input of the change-of-resolution block. The output of the latter is then connected to the input of the decoder, which can be programmed in such a way as to encode the signal at bitrate B2 and then transmit it to the decoder.

[0017] The above solution, which may be defined as an explicit transcoding, presents the fundamental drawback due to the computational complexity linked to its implementation, a computation complexity which renders the solution in question far from attractive in general and in any case unacceptable in those applicational contexts in which a sufficient processing capacity is not available.

[0018] A solution that is alternative to the canonical approach examined previously is described in the European patent applications Nos. 01830084.8 and 01830227.3, both in the name of the present applicant. [0019] The above alternative solution is based upon a "transcoding" process which makes it possible to convert the format, i.e., change the resolution, by operating in the domain of the discrete cosine transform (DCT) of the prediction error. In this way, it is possible to reduce considerably the computational complexity of the operation of format conversion, thus reducing accordingly the implementational complexity.

[0020] In order to achieve a correct anti-aliasing filter-

ing function, it is in any case necessary to resort to linear-filtering techniques, which require, as is known from the theory of digital filtering, the availability of a certain number of coefficients surrounding the ones that are being processed, the aim being to eliminate, or at least reduce as far as possible, edge effects.

Object and summary of the present invention

[0021] The purpose of the present invention is to provide a solution for conversion of format implemented operating in the DCT domain that is able to overcome the intrinsic constraint of the known solutions referred to previously.

[0022] According to the present invention, the above purpose is achieved thanks to a process having the characteristics specified to in the annexed claims.

[0023] The invention also relates to the corresponding system, as well as to the corresponding computer program product which can be loaded directly into the memory of a digital computer and which comprises portions of software code which are adapted to perform the process according to the invention when the computer program product is run on a digital computer.

[0024] The solution according to the invention makes it possible to obtain, in the discrete cosine transform (DCT) domain, block downsampling of an MPEG video bitstream in such a way that the resulting sequence is reduced, for example, to one quarter of the original format

[0025] The solution according to the invention is characterized by its ease of implementation and by the corresponding reduction in computational complexity and in the resulting expenditure of storage space, without this being in any way prejudicial to the quality of the final result.

[0026] The solution according to the invention is therefore proposed as an alternative solution both with respect to frequency conversion performed in the pixel domain by means of FIR pre-filtering and decimation, and with respect to the implementation by means of filtering in the DCT domain.

[0027] In order to achieve the above purpose, obtaining results that are superior or at least equivalent to those of the solutions according to the prior art, the solution according to the invention adopts a non-linear technique for anti-aliasing filtering.

Detailed description of the annexed drawings

[0028] The invention will now be described, purely by way of non-limiting example, with reference to the annexed drawings, in which:

- Figure 1, which regards the prior art, has already been described previously;
- Figures 2 and 3 are schematic illustrations of the organization of the macroblocks and of the corre-

- sponding blocks in the framework of the MPEG standard, as well as of the definition of corresponding microblocks according to the invention;
- Figure 4 illustrates the basic principle of the selection operation performed in the solution according to the invention;
- Figure 5 is a functional block diagram illustrating the processing operations performed in the framework of a system according to the invention;
- Figure 6 illustrates a few examples of masking structures that may be used in the context of the present invention; and
- Figure 7 is essentially a diagram that synthesizes the principles underlying the invention.

Detailed description of an exemplary embodiment of the invention

[0029] As has already been said previously, the conversion of format of a video signal usually involves antialiasing low-pass filtering and subsequent decimation of the filtered pixels.

[0030] Operating in the DCT domain, which is typical of the MPEG standard, the blocks of pixels (usually of size 8x8) are broken down into the equivalent spatial sequences in order to reduce the number of representative coefficients (which may vary according to the bitrate)

[0031] In this way the visibility of the individual pixels is lost, and consequently the low-pass filtering and decimation are carried out on the individual frequencies.

[0032] The variance of the DCT coefficients, and hence their variability, decreases very rapidly with the frequencies. In order to reconstruct an image, it is sufficient to preserve only a relatively small set of terms concentrated around the d.c. coefficient (DC). In addition, it is known that the human eye is less sensitive to high frequencies than to low frequencies, and this implies that minor errors resulting from the elimination of the high frequencies end up not having any particular importance of a subjective nature.

[0033] In the MPEG standard, the 8x8 DCT blocks are grouped together in sets of four in such a way as to give rise to 16x16 macroblocks. Figure 2 illustrates a macroblock MB which groups together four blocks B1, B2, B3 and B4.

[0034] Reducing to one quarter the format of a video bitstream means reducing to one quarter each macroblock, i.e., passing from macroblocks to blocks.

[0035] In the final bitstream, the blocks obtained from downsampling are re-organized into 16x16 macroblocks. The overall number of macroblocks is hence reduced to one quarter between the input bitstream and the output bitstream.

[0036] The solution according to the invention envisages, in its currently preferred embodiment, performing the anti-aliasing filtering by exploiting a non-linear approach that chooses the frequencies of the source im-

20

age that are to be preserved, eliminating the other frequencies. Eliminating a DCT coefficient means removing the energy corresponding to a given spatial sequential component (vertical, horizontal, vertical plus horizontal).

[0037] The low-pass filtering and decimation operations performed are thus organized into two steps.

[0038] The first step basically consists in identifying and preserving the frequencies that are deemed to be significant, eliminating the remaining frequencies. The said selection is carried out within each block B1, B2, B3 and B4 by means of a DCT-coefficient visibility mask applied to a 4x4 "microblock" and setting to zero the coefficients of the other three microblocks.

[0039] Figure 3 represents a generic block Bi, and the microblock selected, which is designated by mBi, is the one illustrated on the top left. The coefficients of the other three microblocks which occupy the positions on the bottom left, top right and bottom right are, instead, set to zero.

[0040] The second step basically consists of a decimation obtained by performing, on the microblock mB previously isolated, the inverse discrete cosine transform (IDCT), which is thus calculated on sixteen coefficients (4x4), and by performing division by an appropriate distortion factor. The IDCT applied to the 4x4 microblock makes it possible to shape further the dark-band spectrum wherever a higher degree of attenuation is encountered. The IDCT is repeated on the microblocks of the four blocks that are to be merged. The four microblocks resulting from the IDCT are gathered together in an 8x8 block to which the DCT is applied.

[0041] The above operating sequence is schematically illustrated in the functional flowchart represented in Figure 4.

[0042] In Figure 4, MB designates the individual macroblock (this will generally be the i-th block in a video sequence) which is to be divided, in a step designated by 10, into four 8x8 blocks designated by B1 B2, B3 and B4

[0043] In a step designated by 12, from each of the blocks B1, B2, B3, B4 a corresponding microblock mB1, mB2, mB3, mB4 is derived. This is done by selecting the 4x4 microblock on the top left — hence the one located in a position close to the d.c. component — according to the modalities described previously, namely setting to zero the coefficients of the other microblocks. [0044] The steps designated by 14 and 16 respectively correspond to a possible masking action performed on each microblock mB1, mB2, mB3, mB4 (this may be done by means of masks, such as the masks M1, M2 or M3 which are illustrated in Figure 6 and which will be described in greater detail hereinafter) and to execution of the IDCT on the microblocks mB1, mB2, mB3, mB4, after possible prior masking.

[0045] The reference number 18 designates the merging of the microblocks thus obtained into a single 8x8 block which, in the step designated by 20, is to un-

dergo the DCT operation with a view to obtaining the final decimated block, designated by MBD.

[0046] The diagram of Figure 5 shows how the aforesaid operations may be implemented by means of appropriate matrix sums and multiplications, remaining in the DCT domain.

[0047] The solution according to the invention is therefore suited to being implemented — according to criteria that are within the reach of a person skilled in the branch on the basis of the detailed description of the invention provided herein — whether by using a specialized processor or by using a general-purpose digital computer (such as a digital signal processor — DSP) on which the computer program product according to the invention is run.

[0048] In particular, Figure 5 considers a single macroblock MB of the four that are usually to be reduced simultaneously. Each 8x8 DCT block B1, B2, B3 and B4 deriving from the macroblock MB is processed in such a way as to preserve only the coefficients corresponding to the microblock mB1, mB2, mB3, mB4 in the top left-hand corner.

[0049] The microblock mB1, mB2, mB3, mB4 is then multiplied by appropriate matrices, on the right and on the left, according to the position of the block. All the matrices considered are of size 8x8 (hence of a rank higher than, namely twice, that of the microblocks to which they are applied) and, in the diagram of Figure 5, they are respectively designated by S tl, D_tl, S_tr, D_tr, S_bl, D_bl, S_br, D_br, where S stands for Sinister (left), D stands for Dexter (right), t stands for top, b stands for bottom, 1 stands for left, and r stands for right.

[0050] Four blocks are thus obtained, designated by B1t, B2t, B3t, B4t (temporary i-th blocks), the sum of which, designated by B1f, corresponds to the merging block of step 18 of Figure 4 which has already undergone DCT (step 20 of Figure 4). In this way it is possible to calculate the IDCT applying the calculation to the 4x4 microblocks (instead of the 8x8 blocks as would normally be the case), then performing the merging and the DCT of the resulting 8x8 block.

[0051] Figure 6, which is divided into three parts designated by a), b) and c), represents three masking structures (with reference of course to a 4x4 microblock) which, once applied to the microblocks on the top left mB1, mB2, mB3, mB4 isolated from the blocks B1 B2, B3, B4, have yielded particularly satisfactory results from the qualitative standpoint. And this, even though the aforesaid masking action entails, in itself, setting to zero the other coefficients of the microblock, namely the coefficients corresponding to the positions appearing in white in the various masks M1, M2 and M3. It will, however, be appreciated that the masks illustrated in Figure 6 are in any case such as to preserve the components that are closest to the d.c. coefficient (DC).

[0052] Of course, the solution described is not limited, as regards its possible applications, to the case alone of microblocks of size corresponding to 4x4 pixels

20

35

40

45

(hence for a reduction to one quarter of the original spatial resolution), but may be generalized to arbitrary sizes of said microblocks.

[0053] In particular, with reference to the schematic representation of Figure 7, it is possible to extract, from each block under examination belonging to a given macroblock (whatever the sizes of said macroblocks and said blocks), a microblock of arbitrary size MxN, provided that this size is smaller than that of the complete block. To each microblock it is then possible to apply the inverse discrete cosine transform IDCT $_{MxN}$, finally proceeding to a reconstruction operation, generically designated by R, where a block of the same size as the original one is reconstructed.

[0054] Of course, without prejudice to the principle of the invention, the details of implementation and the embodiments may vary widely with respect to what is described and illustrated herein, without thereby departing from the scope of the present invention as defined in the annexed claims.

Claims

- A process for converting the format of MPEG video bitstreams, characterized in that it comprises the operations of:
 - identifying DCT macroblocks (MB) comprised in said bitstreams, each macroblock comprising a plurality of blocks (B1, B2, B3, B4);
 - dividing (10) each of said macroblocks (MB) into respective blocks (B1, B2, B3, B4), each block in turn comprising a plurality of microblocks;
 - identifying and preserving, in each DCT block (B1, B2, B3, B4) comprised in a macroblock (MB), the significant frequencies, isolating (12) a corresponding microblock (mB1, mB2, mB3, mB4) of said block (B1, B2, B3, B4);
 - performing (16) an inverse discrete cosine transform (16) on the isolated microblock (mB1, mB2, mB3, mB4);
 - gathering (18) said microblock, which has been isolated and subjected to an inverse discrete cosine transform, together with the homologous microblocks obtained from the other DCT blocks comprised in a respective macroblock (MB), so as to obtain a merging block; and
 - applying to said merging block the discrete cosine transform (20) so as to obtain a final block that can be assembled into a macroblock (MBD) with converted format.
- The process according to Claim 1, characterized in that it comprises the operation of identifying said DCT macroblocks (MB) as 16x16 macroblocks.

- The process according to Claim 1 or Claim 2, characterized in that it comprises the operation of dividing said macroblocks (MB) into 8x8 DCT blocks (B1, B2, B3, B4).
- 4. The process according to any one of the preceding claims, characterized in that it comprises the operation of selecting said corresponding microblock (mB1, mB2, mB3, mB4) as a 4x4 microblock.
- 5. The process according to any one of the preceding claims, characterized in that said corresponding microblock (mB1, mB2, mB3, mB4) is chosen in the area around the d.c. component of the respective block (B1 B2, B3, B4).
- 6. The process according to any one of the preceding claims, characterized in that said corresponding microblock (mB1, mB2, mB3, mB4) is identified as the microblock on the top left-hand side of the respective block (B1, B2, B3, B4).
- 7. The process according to any one of the preceding claims, characterized in that said operations of performing, on the isolated microblock (mB1, mB2, mB3, mB4), an inverse cosine discrete transform (16), gathering (18) said microblock, which has been isolated and subjected to an inverse discrete cosine transform, with the homologous microblocks obtained from the other DCT blocks comprised in a respective macroblock (MB), and applying to said merging block the discrete cosine transform (20) are carried out multiplying said isolated microblock and said homologous microblocks (mB1, mB2, mB3, mB4) by respective multiplication matrices (S_tl, D_tl, S_tr, D_tr, S_bl, D_bl, S_br, D_br) so as to obtain respective temporary blocks (B1t, B2t, B3t, B4t) and adding together said respective temporary blocks (B1t, B2t, B3t, B4t), the sum of said temporary blocks identifying said final block that can be assembled into a macroblock (MBD) with converted format.
- 8. The process according to Claim 8, characterized in that said multiplication matrices (S_tl, D_tl, S_tr, D_tr, S_bl, D_bl, S_br, D_br) are of a rank higher than, and preferably twice, that of said isolated microblock and said homologous microblocks (mB1, mB2, mB3, mB4).
- 9. The process according to any one of the preceding claims, characterized in that it comprises the operation of subjecting said corresponding microblock (mB1, mB2, mB3, mB4) to masking, using a mask (M1, M2, M3) that acts on a subset of the microblock (mB1, mB2, mB3, mB4) itself.
- 10. The process according to any one of the preceding

55

claims, **characterized in that** said operation of isolating said corresponding microblock (mB1, mB2, mB3, mB4) involves setting to zero the coefficients associated to the microblocks of said block (B1, B2, B3, B4) other than said corresponding microblock (mB1, mB2, mB3, mB4).

11. The process according to any one of the preceding claims, **characterized in that** said inverse discrete cosine transform (IDCT) is calculated on sixteen coefficients.

12. The process according to any one of the preceding claims, **characterized in that** it comprises the operation of applying a distortion factor to the result of said inverse discrete cosine transform (IDCT).

13. A system for converting the format of an MPEG video bitstream, in the form of a dedicated computer circuit operating according to the process as per any one of Claims 1 to 12.

14. A computer program product directly loadable into the memory of a digital computer and comprising software code portions for performing the operations according to any one of Claims 1 to 12 when said product is run on a computer.

30

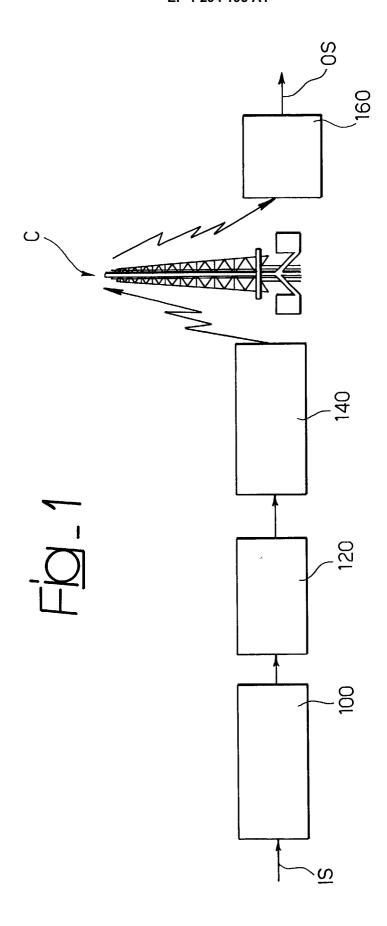
35

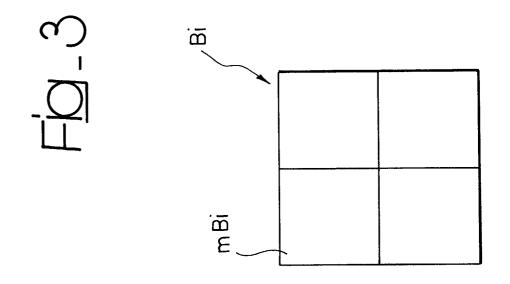
40

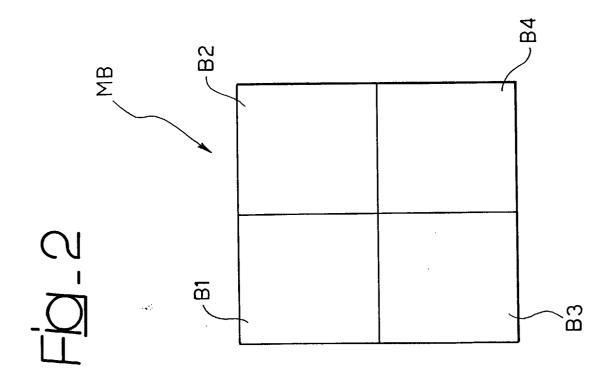
45

50

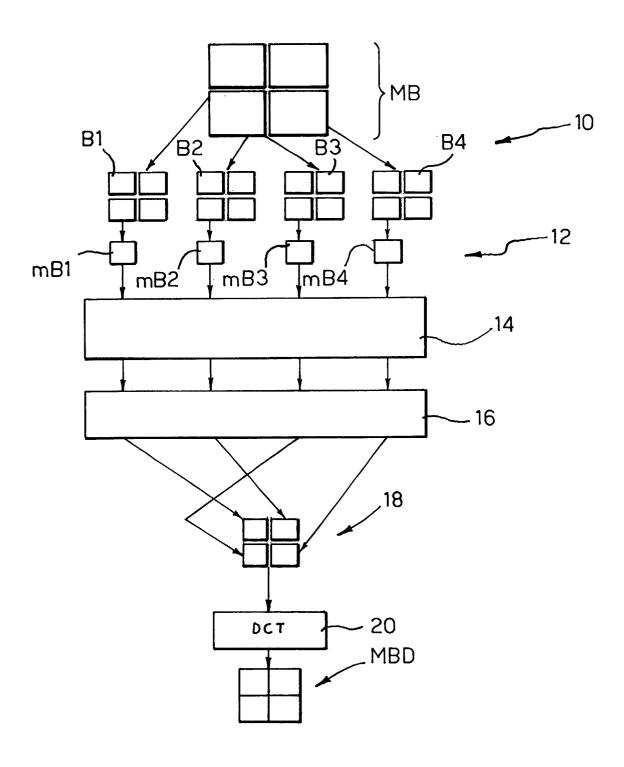
55



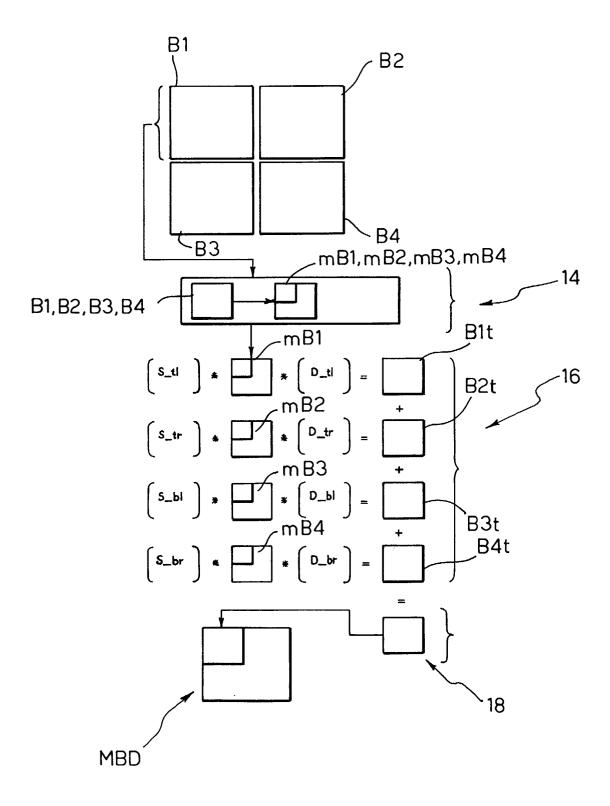




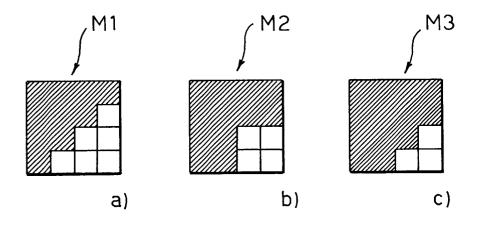


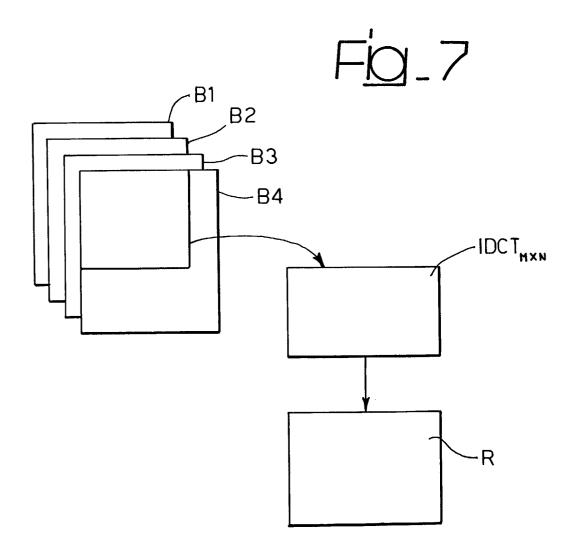














EUROPEAN SEARCH REPORT

Application Number EP 01 83 0589

Category	Citation of document with in of relevant passa	dication, where appropriate, ges	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.CI.7)	
X	size change in the IEEE TRANSACTIONS O	N CIRCUITS AND SYSTEMS Y, APRIL 2001, IEEE, ges 461-474,	1-14	H04N7/26 G06T9/00	
X	INVERSE MOTION COMP IEEE TRANSACTIONS O FOR VIDEO TECHNOLOG US,	WN-SAMPLING AND FOR ENSATION" N CIRCUITS AND SYSTEMS Y, IEEE INC. NEW YORK, ne 1997 (1997-06-01), 0690585	1-14		
Α	WO 00 51357 A (SARN 31 August 2000 (200 * page 7, line 25 - figure 1 *	0-08-31)	1-14	TECHNICAL FIELDS SEARCHED (Int.CI.7) H04N G06T	
A	IN THE DCT DOMAIN" PROCEEDINGS OF THE VA, US,	SCODING DV INTO MPEG-2 SPIE, SPIE, BELLINGHAM, 1998, pages 1026-1032, */	1-14		
	The present search report has b	peen drawn up for all claims			
	Place of search	Date of completion of the search		Examiner	
	THE HAGUE	6 March 2002	Geo	orgiou, G	
X : part Y : part docu A : tech O : non	ATEGORY OF CITED DOCUMENTS icularly relevant if taken alone icularly relevant if combined with anothem of the same category inclogical background rewritten disclosure imediate document	L : document cited fo	ument, but publi e n the application or other reasons	ished on, or	

EPO FORM 1503 03.82 (P04C01)



EUROPEAN SEARCH REPORT

Application Number EP 01 83 0589

Category	Citation of document with ir of relevant passa	idication, where appropriate, ages	Relevant to claim	CLASSIFICATION APPLICATION (I	
Ą	WO 00 00907 A (SKOD	RAS ATHANASIOS ILAOS (SE); ERICSSON 2000 (2000-01-06)	1-14	APPLICATION (I	nt.Cl.7)
				TECHNICAL FIEL SEARCHED	DS (Int.CI.7)
	The present search report has t	peen drawn up for all claims			
	Place of search	Date of completion of the search		Examiner	
	THE HAGUE	6 March 2002	Geo	rgiou, G	
X : part Y : part docu A : tech O : non	ATEGORY OF CITED DOCUMENTS icularly relevant if taken alone icularly relevant if combined with another iment of the same category inclogical background written disclosure mediate document	E : earlier patent after the filing ner D : document cite L : document cite	d in the application d for other reasons	shed on, or	

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO.

EP 01 83 0589

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

06-03-2002

	Patent document cited in search report		Publication date		Patent famil member(s)		Publication date
WO	0051357	A	31-08-2000	EP EP WO WO	1155573 1118224 0051357 0019725	A1 A1	21-11-2001 25-07-2001 31-08-2000 06-04-2000
WO	0000907	A	06-01-2000	AU EP SE WO US	4940399 1090359 9802286 0000907 2001006561	A1 A A1	17-01-2000 11-04-2001 27-12-1999 06-01-2000 05-07-2001
	is then over two two was now that dath over deer		. an on on wa 80 MM MM (S MP N) We 64 NP N	nd ann add unt mus a	en en sue eus une uné sur mié des des die but un	allo and and and the the the the the	man elde kån kan den men apn kån sam som som

 $\frac{Q}{M}$ For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

FORM P0459